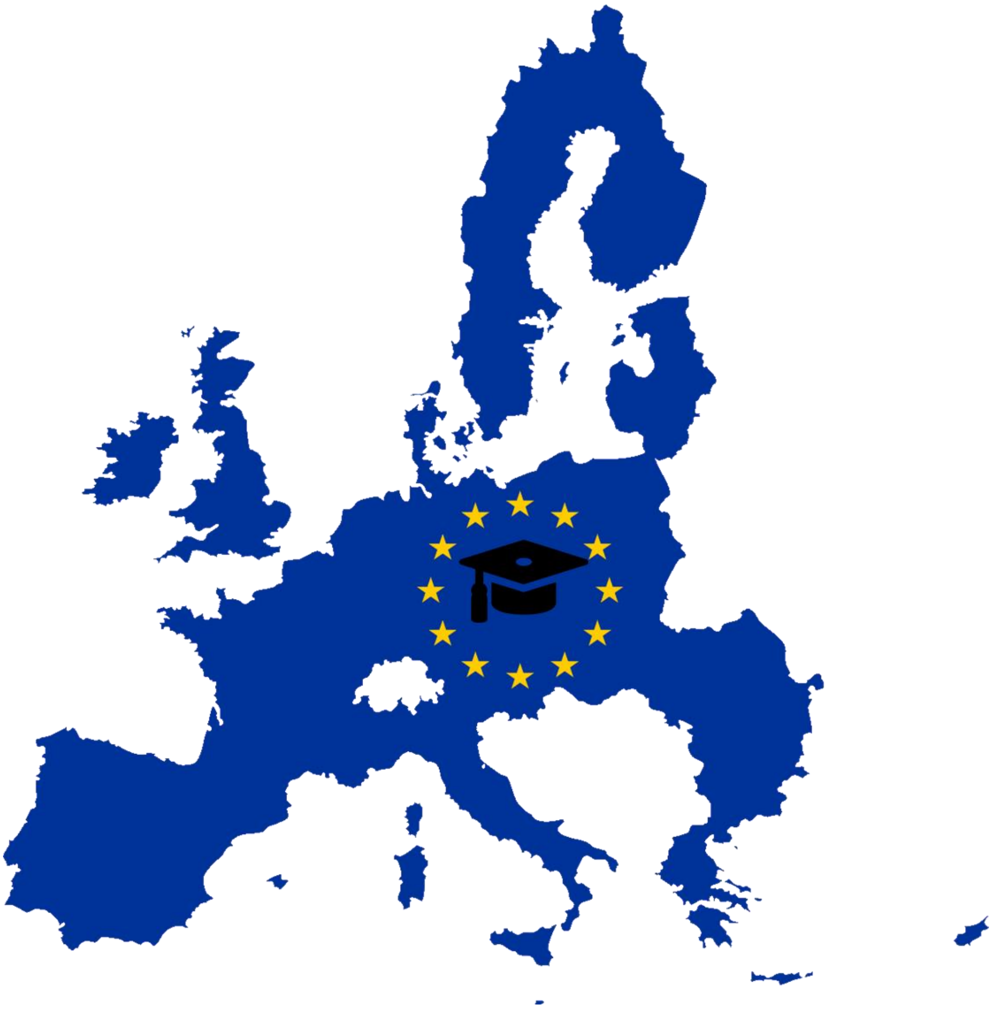
SCIENCE PATTERNS IN THE EUROPEAN UNION



Bachelor Thesis
disciplinary thesis of the study Liberal Arts and Sciences
Major: Natural Sciences & Innovation Management

Date of submission: **26th of January 2018**
Wordcount: **6750 words**

Author: **A.E. (Anne) Jonker, 4092767**
Supervisor: **dr. G. J. Heimeriks**

Utrecht University

Abstract

In the Europe 2020 Strategy is stated that the European Union wants to be a world leader when it comes to innovation. Considering the innovation performance of all EU Member States, there is still a long way to go before this goal is met. As innovation and the generation of knowledge are strongly related. A better understanding of the knowledge generation of EU countries may pave the way to meet the goal. No such study is performed in recent years. In this descriptive study, quantitively the development of the science systems of all 28th Member States are analysed. The findings are: a significant occurrence of convergence among the countries. And the conclusion that sciences of main importance across the EU have extended in comparison to findings of earlier studies. These new science domains do match focus topics of the Horizon 2020 research funding programme.

Acknowledgements

In advance of writing my bachelor thesis, many years of study have been passed. In the elapsed time, even more people have been important to me and make the making of this closing piece possible. Many thanks to all of them, without your teaching, advice, chances I got, or support, my study and either this would not have been the same.

However, reflecting on the thesis process itself, I would like to thank some people in particular. First of all, my supervisor dr. Gaston Heimeriks. It was a pleasure to have him as an understanding and supportive supervisor. I also would like to thank Denise Reike for her support and advice during the process, as well for the tip to contact Elma Zijdema. I have to thank her for the insightful conversations and the advice. Dr. Iris van der Tuin, the director of my study Liberal Arts and Sciences, to her a warm hearth for her trust and flexibility to make this project work and worth in context. Of course, a word of thank to the men who were in the same working group as I, their feedback have been a good contribution.

Next, I would like to thank some dearest friends; Pankaj and Nina. For their offers to help, and the happy, off-topic chats which were very welcome during the past ten weeks.

Lastly, I would like to thank my great family for their support and understanding that this work was my main priority. Especially I would like to thank prof. dr. Frans Oort, for playing a major role throughout my studies, and in context with this project for the fruitful feedback, the sharpening questions, and the offers for help when needed. Of course, many thanks to my parents for their unconditional support and advice. Thanks to my dear brother, for his encouragement, and in time of stress I would like to be more like you.

I would like to thank you, dear reader, for taking the time and having the interest to read my work. I wish you a joyful read. Any points of discussion or recommendations can be send to jonker_anne@hotmail.com.

Table of Contents

Abstract 2

Acknowledgements 2

Introduction 4

Theory 6

- Path dependency of knowledge development 6
- Convergence & Convergence Clubs 7

Methodology 9

- Hypothesis 1 & 2 9
- Hypothesis 3 & 4 10

Results 11

- Hypothesis 1 & 2 11
- Hypothesis 3 & 4 13

Discussion 16

- Findings 16
- The research 17

Conclusion 18

Resources 19

The data collection that was used for the quantitative analysis, can be submitted by mail on request.

Introduction

The history of the European Union is extensive and dynamic. The treaty of Rome (1959) is considered as the start of the European Union (Hashimoto & Rhimes, 2017). This was the first treaty of the European Economic Community (ECC) and included as most typical agreement the free movement of goods, services and labour across the contracted nations. The main intention of this alliance was that it would improve the cross-national political relations and increase the economic position of each country through the interconnectedness (Dedman, 2006; Warleigh-Lack, 2008). Over the years, six more countries accessed the ECC. In 1993, the ECC consisted by then of twelve countries, the European Union was found as it is today[1]. In the following years, the EU expanded to a highly interconnected institution of twenty-eight Member States. These countries share political stability, free movement of goods, services, and people across accessioned states, and a shared goal of "economic growth" within each nation (Dedman, 2006).

The increase of economic prosperity is strongly associated with the innovativeness of a system (Hirooka, 2006). In this study a "system" refers to a country. In order to innovate, new insights are required (Popadiuk & Choo, 2006). Knowledge generation is important to establish these insights. Summarising this argumentation, knowledge generation could be considered as an indicator of innovativeness. This argument is supported by many scholars (e.g. Freeman & Soete, 2009; Quintane, Casselman, Reiche, & Nylund , 2011).

The relation between innovativeness and knowledge generation is also taken into account by the EU (European Commission (EC), n.d.a) Since 1981 the Single European Act, the national policy guidance for Member States, covers the importance of research stimulation at national level (Reillon, 2015). In the current guidelines for a future European Union, the *Europe 2020 Strategy*, covers research and innovation by a special programme (EC, n.d.b). The *Horizon 2020* is a funding programme that facilitates in particular research and development as driver of innovation. Implicitly, the Horizon 2020 programme is not the only transnational funding programme, but many agents consider it as the most important programme to meet a goal of the future European Union; being a world leader when it comes to innovation (Veugelers, 2016)

Unless, many European countries are listed as "top innovators" (Weller, 2017), the rapport by Veugelers (2016) draw a different conclusion when considering the "European Union". The rapport states that especially the *modest innovators* in the EU have a hard time to catch-up with even the *moderate innovators*. Various causes are mentioned, but so far no clear reason can be designated. As innovation is in many cases the result or a side product of science (Audretsch, 2008). The question rises if the lagging countries are not able to produce sufficient knowledge. Bonaccorsi (2007) provided a little insight on the knowledge generation of countries in the EU. However, the study only focused on three science subdomains; computer sciences, high energy physics, and nanotechnology, and the dataset for this comparative study covers the period from 1990 until 2001. An extensive search for more recent studies on this topic, gave no results. At global level, Horlings & van den Besselaar (2011) investigated the generation of knowledge and determined a classification of nations according to their scientific focus. The division is classified by convergence clubs. Among science systems that are part of a convergence club, convergence takes place with a focus on typical science domains. Convergence is a process whereby smaller science systems grow faster than larger, more developed science systems. The current Member States are divided over different convergence clubs by the aforementioned study. This indicates that the countries in the European Union are not congruent in their science systems, although they should perform in accordance with the "bigger system" they belong to; the "European Union" convergence club. Considering the fact that the science systems are dynamic (Heimeriks & Boschma, 2014), the classification of countries in the European Union may have changed since 2008.

[1] The first treaty of the European Union is the treaty of Maastricht. In 2007 the treaty of Lisbon was signed by all Member States and came into effect in 2009. It amends previous treaties and forms with them consolidation of the European Union of today (Hashimoto & Rhimes, 2017).

Therefore, it is considered as relevant to perform a similar study on the occurrence of convergence. To make a start on the further understanding of the knowledge generation and its development in the EU, the following research question is addressed:

“What convergence patterns in science systems can be identified among the Member States of the European Union?”

The decision for the plural “convergence patters” refers to the twofold of the answer. This study analyses the occurrence of convergence among the Member States of the EU. Furthermore, the appearance of the “European Union” convergence club is investigated. After addressing these aspects, the research question can be answered in accordance to the scope of this research.

As previously stated, the findings will make an academic contribution because of its more state of the art character compared to other studies on this topic. The analyses will show at what pace and within which area scientific development of the different countries in the EU takes place. Due to the descriptive aim of the research, the outcome could be used as a starting point for more explanatory studies on this topic. The conclusion of this study will be valuable for the European Commission too. It provides insight into the generation of knowledge and most relevant science domains among the Member States of the European Union. It is important to production of science is one of the indicators of the innovativeness of a system. Therefore, the European Commission can consider the results when plans are made for more research funding programmes.

To execute this study, which is about convergence and convergence clubs, the understanding of the path dependency is assumed as necessary. Therefore, in the next section this theory is shortly introduced and relevant concepts are further clarified. In the second part of the theory, convergence and convergence clubs are explained and the hypotheses are formulated. In the subsequent section, the methods of this research are pointed out. To address the research question a quantitative study will be performed. In the following chapter the results are presented and analysed. The convergence pattern of the scientific output will be mapped and quantified for all 28 EU Member States. Followed by the results of testing the hypotheses to identify the convergence club(s) among the Member States. The shortcomings and implications of the findings are covered in the discussion section. This thesis is ended by a conclusion. The final section will also provide some advice considering the science system and related incentives for its development within the European Union.

Theory

This study is about the twenty-eight Member States of the European Union. Despite the great degree of solidarity and interconnectedness among the nations, each natation is sovereign and has demarcated borders (Belammy, 2017). For this reason, each country is regarded as a system. According to concepts of the path dependence theory and how it affects the development of a system, it is reasonable to classify science systems by the boundaries of a country (Horlings & van den Besselaar, 2011; Boschma & Capone, 2016). A science system is considered is a subsystem of a nation (Katz, 1999) wherein many types of agents interact for the sake of knowledge progression (Kessels, 2013). In the next subsection relation between the path dependence theory and the development of science systems is elaborated upon. This understanding is relevant because it shows how convergence patterns and convergence clubs developed. This is the topic of the last subsection of the theory and the hypotheses derived from it are also presented here.

Path dependency of knowledge development

The path dependence theory comes from the economic discipline, but is applied in many studies by other sciences (Hirsch & Gillespie, 2001). Path dependency of a country reflects the accessibility to its resources like natural resources, human capital, and monetary capital (Boschma & Capone, 2015). A country's history, and the events which took place affect the development of a nation by restraining future choices, the possible "paths" will be determined (Valdaliso et al., 2014). However, path dependency is also a co-evolving process, whereby new events reshape future opportunities (Boschma & Capone, 2016). Furthermore, political and cultural history shapes the countries development and future possibilities for development (Mahoney, 2000; Valdaliso et al., 2014; Boschma & Capone, 2016).

According to the path dependence theory, the individual development of the Member States of the European Union is shaped by the accession process. To become a member of the EU, a country has to meet Convergence Criteria of the European Union (Hashimoto & Rhimes, 2017). These criteria reflect the economic situation of nation. In addition, all Member States must recognize and accept fundamental legislation (*acquis communitaire*) of the European Union (Hashimoto & Rhimes, 2017). Due to the equity principals of resources and legislation and the share of common goals, the assumption is that Member States are subjected to a similar pattern of path dependency.

Although exact relations are un clear, national path dependency "guide" the generation of knowledge (Horlings & van den Besselaar, 2011). Therefore, the accession of a nation to the EU is considered as an aspect that influence the development of a national science system within a certain direction. Nonetheless, Heimeriks & Boschma (2014) state that previous generated knowledge is a better indicator to forecast the development of a science system.[2] Bearing in mind this argument, a difference would be expected between the science systems of the different countries in the European Union. On the other hand, after accession the interrelatedness of the countries increases, which argues for a similarity among the different science systems . Enforced by changes of the access to resources through for example the free movement of labour and goods (Boschma & Capone, 2016) and the share of common research funds and goals (Heimeriks & Boschma, 2014; Horizon 2020, 2018; Hashimoto & Rhimes, 2017). The pace of a change of a science system depends on the relatedness to other systems (Boschma & Capone, 2016). Thus, the moment of accession by a country is considered as relevant. The statement that science systems are dynamic (Heimeriks & Boschma, 2014) supports the idea that a shift among the different science systems over time is reasonable. How this shift take place and what is expected for the development of Member States' science systems is explained in the next subsection.

[2] This concept is also known as the "adjacent possible" (Kauffman, 1993).

Convergence & Convergence Clubs

Another theory from the economic discipline is that similar systems, like industries in a shared sector, grow to a common equilibrium, the steady state, shaped by the context within they develop (Curran & Lekker, 2011). As in the previous subsection is shown, when a country has joined the European Union, simply said, the country becomes part of a "bigger system". The growth to the steady state is captured by the economic theory of "convergence" (Solow, 1958) which refers to phenomena of typical growth patterns of related systems that are in development. Briefly, the theory states that systems which are further removed from the shared "steady state equilibrium", approach this point per development-ratio faster than systems that are closer. The concept can be explained by path dependence and is applicable on science systems (Horlings & van den Besselaar, 2011). In case of science systems, it means that smaller science systems grow faster than larger science systems in terms of their scientific output. The limit of growth, the steady state of science systems, is not clear. The occurrence of convergence in a science system points out that the system is in development (Horlings & van den Besselaar, 2011).

Considering the increase of interrelatedness of countries due to their accession to the EU and the occurrence of convergence among related science systems. The assumption of the author is that convergence has taken place between the science systems of Member States (H1). Implicitly, the following condition must hold: the size of the considered science systems differs at the initial moment of measurement. If this condition is true, the combination of the first hypotheses with the path dependence theory in respect to different moments of accession leads to next hypothesis: the pattern of convergence is stronger since 2004 (H2). The underlying thought is that in 2004 the largest enlargement at once took place.[3]

Another conclusion by Horlings & van den Besselaar (2011) is the existence of "convergence clubs". Convergence in science seem to happen within a group of countries of a same dominator, like ecological strongholds, or in this study countries that are in the European Union. The division can be understood by the theory of path dependence; the national context shapes the possibilities and opportunities of a science system. The identified convergence clubs by Horlings & van den Besselaar (2011) also point at relevance of path dependency shaped by the original science portfolio of a nation. The countries that have belonged to other "bigger systems" like former soviet countries which are now developing countries, appear in the same convergence club. The division made by the authors is presented in table 1 and the accompanying geographical distribution. Regarding the distribution, also the economic development suggests to be important. For the consistency of the study the division is based on the moment of accession. The first group consists of the EU-15 countries (all listed as "High-income industrialized nations" in Table 1). The second group are the other thirteen Member States. This group is considered as whole because the three countries that joined the European after 2004 show up in the same convergence clubs where to other new Member States belong according to Horlings & van den Besselaar (2011).

[3] The "old" Member States are the first twelve and taking into account the large time span between the fourth enlargement (accession of Finland, Sweden and Austria in 1995) and the fifth enlargement (2004) combined with adaptability of science systems subjected to interrelatedness that increase over time, these countries are also considered as "old" Member States. The "new" countries are the countries that accessioned the European Union since 1995.

Convergence club name	Countries characteristic of the factor	Main scientific specialisation
High-income industrialized nations	Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, United Kingdom (Cyprus & Malta)	Clinical medicine and basic medicine; biological sciences
Emerging economies	Croatia, Czech Republic, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, Slovenia	Computer science; physical sciences; electrical engineering and other domains of engineering
Former Soviet Republics	Bulgaria, Estonia (& Latvia, Lithuania?)	Chemical sciences; physical sciences; mathematics; materials engineering and electrical engineering

Table 1: different convergence clubs (Horlings & van den Besselaar (2011))[4].

Following the argumentation that joining the European Union affect the science system of a nation. As a result of “new” interrelatedness among countries and the co-evolvement of path dependency that is assumed to happen in science systems. The expectation is that science portfolios of Member States of the European Union become more similar over time and converge to a common convergence club (H3). As most countries belong to the “High-income industrialised nations” convergence club according to Horlings & van den Besselaar (2011), the assumption is that the main focus of the science output will be in the domains of medical and biological sciences. However, interrelatedness also enforces the transmission from knowledge among actors (Boschma & Capone, 2016). Therefore, the expectation is that the “European Union” science systems in the final observations consist of more focus science domains than the science system in the initial situation (H4).

[4] In case a country was not listed, the website of the International Monetary Fund (IMF, 2018) has been consulted to identify the economic position of the country. This is applied on Cyprus and Malta, both are identified by the IMF as advanced economies, similar to other EU-15 countries. Based on the current economic situation of Finland, this country is also listed in the group of *High-income industrialized nations*. However, Horlings & van den Besselaar (2011), gathered Finland under *Emerging Economies*. Horlings & van den Besselaar (2011), list Latvia and Lithuania under *Emerging economies*, however this are also fomer countries in the Soviet Union.

Methodology

To address the research question, a quantitative analysis (Bryman, 2012) is performed to test the different hypotheses presented in the theory section. To achieve saturation in the theory section and conceptualize the different variables, search engines *Google Scholar* and *Scopus* have been used to obtain academic literature. The search process has been expanded by knowledge provide on the website of the European Union (European Union, 2017).

The data for the quantitative analysis was retrieved from “Scimago Journal & Country Ranking” (SJR, 2016). An extensive dataset considering 239 countries, their scientific output for every year from 1996 and 2016, 27 major thematic categories, and 313 specific subject categories, is the source of the relevant data for this research. Although the original dataset covered several categories related to the scientific output of the nations, only the publications are taken into account. This is the same method as Horlings & van den Besselaar (2011) applied in their study. For every Member State (nominal variable), the citable documents produced in every specific category over the full time range (21 years) are included in the prior data selection. The output per specific category is for the further data processing accumulate under dominator of the thematic category, resulting in the interval variable “year_category”. These data were exported to Excel (Microsoft, 2016) to be further processed. In case no output was registered in a “year_category”, the “missing value” (after verifying with the original dataset) had been manually replaced by a zero. This final data matrix is the base for every analysis performed to test the hypotheses.

Considering the fact that the data was retrieved from an independent data source which relies on the Scopus® database of publications, makes the data very reliable (SJR, 2016)[5]. Also the data preparation is done with the greatest precision and with the use of programmes offered by Microsoft Office (Access, Excel, 2010 & 2016). All following steps and tests for the analysis are performed in Excel, using the “Analysis ToolPak”. Because of the specific focus of the research, the external validity is relatively small (Bryman, 2012). As the study only considers the scientific output of nations, the internal validity is hard to justify as possible variables that may affect the outcomes are not taken into account.

Hypothesis 1 & 2

To test hypotheses 1 and 2 the adjusted function by Sala-i-Martin (1996) is appropriate to use (Horlings & van den Besselaar). The outcome of the function determines the occurrence of *absolute convergence*: convergence that is estimated without taking into consideration any sight effects like population growth. The equation that describes the function is as followed:

$$\phi_{t,t+T} = \beta \log(y_t) + \alpha + \varepsilon_t,$$

The left-hand side of the function is the growth rate of the national scientific output of country “y” over a time span t – t+T (in years, t = the initial year, T + t is T years later). The β on the right-hand is the indicator of convergence, in case convergence occurs in a system β has a negative value. Whereby a more negative value means stronger convergence among science systems (Horlings & van den Besselaar, 2011). The other values indicate the intercept (α), the total scientific output of country “y” at time t, and the error term of nation “y” with respect to the regression function. To test for convergence the regression line have been requested in a scatter plot including all 28th “$\phi_{t,t+T}$” (Y-axis and on the X-axis log(y_t)). For every “$y_{t(+T)}$” the entire scientific output of a nation is considered in the given year. This value is calculated by summing all values of the 27 science categories. After this, to calculate all the “$\phi_{t,t+T}$” the following equation have been entered in Excel:

=(log([scientific output country X at time t+T]/ ([scientific output country X at time t]))/T

For hypothesis 1 according to the year of accession of most member states, the values for time are t = 2004 and t+T = 2016 (T = 12). For both points in time the total scientific output corresponding to these years per country (all 28th Member States) are entered in the formula.

To test hypothesis 2 the value of hypothesis 1 is compared to the value calculated when t = 1996 and t+T = 2004 (T = 7). To test whether convergence has increased since 2004, the β-values are

[5] A full description of the methods used by Scimago is available at: http://www.scimagojr.com/aboutus.php

compared. If convergence have become stronger since the biggest enlargement of the European Union in 2004, the following must be true: $\beta_{04\text{-}16} < \beta_{96\text{-}04}$

Hypothesis 3 & 4

Convergence clubs are identified by measuring the intensity of output in the different science domains (Horlings & van den Besselaar, 2011). Intensity can be described with the following equation (derived and defined from the aforementioned article):

$$Intensity_{science\ domain\ x, country\ y, year\ z} = \frac{Scientific\ Output_{science\ domain\ x, country\ y, year\ z}}{\sum Scientific\ Output_{country\ y, year\ z}}$$

The equation describes the distribution of activity in the different science domains by a country for a certain year. The function was applied on the dataset (year_category). Resulting in a relative scientific output for each Member State per science domain for the considered years.

When hypothesis 3 is true, the dispersion of intensity among the countries should be smaller in the final year of measurement when comparing to an earlier point in time. 2004 Is chosen as the earlier point in time because in that year the big enlargement with ten countries took place, Since, only three more countries joined the European Union, all 28th Member States are included for both years. The final year is 2016. To test hypothesis 3 an ANOVA is performed as it shows what the differences are between and within groups.

The values of variance within groups will give an insight if hypothesis 3 does not need to be rejected instantly. If this number is smaller for the measurement of the dataset of 2016 compared to 2004, it means that the countries show a similar dispersion of intensity when considering all science domains. In other words countries have shift their relative focus to shared science domains.

The comparison of the values of variance between groups for the two points in time reflects the diversity of intensity among the science domains. This value cannot be interpreted individually, but need more context. The method that is used to test hypothesis 4 gives a part of the required context. The decision is made to base the argumentation for direct rejection or acceptation within the scope of this research for hypothesis 4 solely on an interpretation of two graphs. The first graph presents the intensity per science domain for each country in 2004 and 2016. The second graph shows the cumulated output per science domain (the sum of the output by all countries in a category) for the year 2004 and 2016. For both years also the output per set of countries is given. As earlier discussed, a distinction is made between the EU-15 countries and the other thirteen countries (new Member States (NMS)).

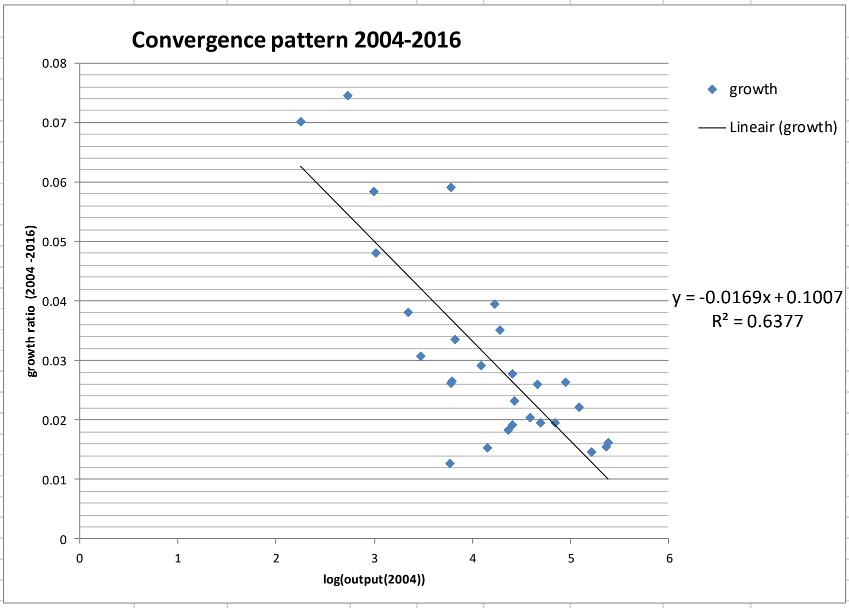
Results

In the previous two sections the four hypotheses which are relevant to answer the main research question are introduced. In this section the results of the different tests are presented and briefly discussed. A more contextualised interpretation is part of the Discussion.

Hypothesis 1 & 2

The first hypothesis states that the convergence among the science systems of the different Member States have occurred. To test this statement the convergence function is plotted for the scientific output by all countries in 2004 and 2016 (in logarithmic scale). Below the output of the analyses performed as described in the methodology. On the x-axis the logarithmic of the scientific output for the year 2004 per country. On the y-axis the calculation of the growth rate of scientific output between 2004 and 2016.

Country	log (output(2004))	growth ratio
Austria	4.427	0.023
Belgium	4.586	0.020
Bulgaria	3.771	0.013
Croatia	3.778	0.026
Cyprys	2.991	0.058
Czech Republic	4.279	0.035
Denmark	4.404	0.028
Estonia	3.340	0.038
Finland	4.409	0.019
France	5.216	0.015
Germany	5.361	0.016
Greece	4.359	0.018
Hungary	4.152	0.015
Ireland	4.088	0.029
Italy	5.081	0.022
Latvia	3.013	0.048
Lithuania	3.474	0.031
Luxembourg	2.731	0.075
Malta	2.255	0.070
Netherlands	4.842	0.020
Poland	4.660	0.026
Portugal	4.227	0.040
Romania	3.775	0.059
Slovakia	3.826	0.033
Slovenia	3.785	0.027
Spain	4.942	0.026
Sweden	4.692	0.020
United Kingdom	5.377	0.016



SUMMARY OUTPUT (2004 - 2016)

Regression Statistics	
Multiple R	0.79855657
R Square	**0.6376926**
Adjusted R Square	0.62375770
Standard Error	0.01031447
Observations	28

ANOVA

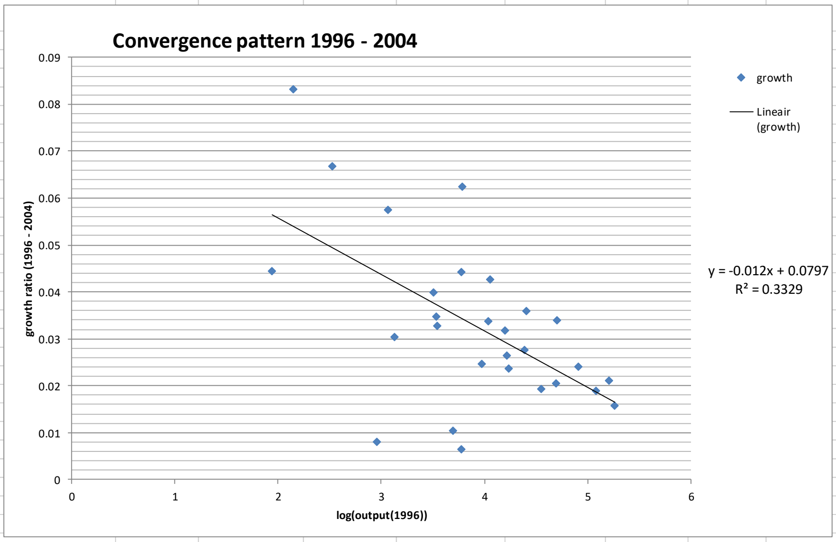
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.004868571	0.004868571	45.76226513	3.52812E-07
Residual	26	0.002766096	0.000106388		
Total	27	0.007634667			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	0.100694246	0.010497909	9.591838031	5.01262E-10	0.079115484	0.122273008	0.079115484	0.122273008
log(2004)	**-0.01686693**	0.002493345	-6.764781233	3.52812E-07	-0.021992079	-0.01174179	-0.021992079	-0.01174179

Figure 1: Results of the test to validate hypothesis 1.

The results presented in Figure 1 clearly confirm hypotheses one. The regression function that describes the relation between the total scientific output in 2004 and the growth ratio over 2004-2016 for all 28 countries has negative value for β (-0.0169). This proves the occurrence of convergence between the science systems for the considered period. The level of R^2 (0.65) indicates that the convergence pattern can be quite good predicted when regarding the initial value for scientific output of a country. Both coefficients of the function have a valid (<0.05) p-value.

Country	log (output(1996))	growth ratio
Austria	4.204	0.032
Belgium	4.393	0.028
Bulgaria	3.699	0.010
Croatia	3.549	0.033
Cyprys	2.524	0.067
Czech Republic	4.043	0.034
Denmark	4.238	0.024
Estonia	3.128	0.030
Finland	4.224	0.026
France	5.084	0.019
Germany	5.214	0.021
Greece	4.061	0.043
Hungary	3.980	0.025
Ireland	3.779	0.044
Italy	4.913	0.024
Latvia	2.957	0.008
Lithuania	3.072	0.057
Luxembourg	2.149	0.083
Malta	1.944	0.044
Netherlands	4.699	0.020
Poland	4.409	0.036
Portugal	3.790	0.062
Romania	3.532	0.035
Slovakia	3.781	0.006
Slovenia	3.506	0.040
Spain	4.705	0.034
Sweden	4.557	0.019
United Kingdom	5.266	0.016



SUMMARY OUTPUT (1996 - 2004)

Regression Statistics	
Multiple R	0.57698531
R Square	**0.33291205**
Adjusted R Square	0.30725482
Standard Error	0.01491455
Observations	28

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.002886292	0.002886292	12.97537061	0.001307404
Residual	26	0.005783542	0.000222444		
Total	27	0.008669834			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	0.079719718	0.013313876	5.987716775	2.54121E-06	0.052352654	0.107086782	0.052352654	0.107086782
log(1996)	**-0.01200**	0.003330275	-3.602134175	0.001307404	-0.018841578	-0.00515062	-0.018841578	-0.00515062

Figure 2: The output for testing the presence of convergence between 1996 – 2004.

The occurrence of convergence between 1996 and 2006 is less strong (Figure 2). According to standards, the prerecession of the function is weak (R-value of 0.333). However, the p-values for the coefficients are acceptable. According to the β-value (-0.0120), convergence among the science systems happened the considered time. When this value is compared to the value found in the previous test, hypothesis two is confirmed when solely estimating the β-values ($\beta_{04-16} < \beta_{96-04}$): -0.0169 ($\beta_{04-16}$) < -0.0120 ($\beta_{96-04}$). With use of the "FISCHER" function in Excel, the z-values and the comparative z-value of the correlation coefficients of both functions are calculated (Statistic Solutions, n.d.) and the p-value of 0.122831 is determined. This implies that hypothesis 2 is not statistically valid.

Hypothesis 3 & 4

Below the results of the test performed to validate hypothesis 3 as described in the methodology. To give the values estimated through the ANOVA context, in the most right column of each annual output the total output for each science domain is listed as well.

SUMMARY	Anova: Single Factor			2004	
Groups	*Count*	*Sum*	*Average*	*Variance*	*(Total output)*
Medicine(all)	28	4,689852196	0,167494721	0,003287098	234348
Biochemistry, Genetics and Molecular	28	2,595274261	0,092688366	0,000676127	129519
Engineering(all)	28	2,920321602	0,1042972	0,000661671	119513
Physics and Astronomy(all)	28	2,705865081	0,096638039	0,001634179	114442
Materials Science(all)	28	2,335776048	0,083420573	0,001609841	93388
Chemistry(all)	28	1,713152244	0,061184009	0,000659738	72123
Computer Science(all)	28	1,448163941	0,051720141	0,000339293	61476
Mathematics(all)	28	1,473478192	0,052624221	0,000355557	58259
Agricultural and Biological Sciences(all)	28	1,380000654	0,049285738	0,000222016	56793
Earth and Planetary Sciences(all)	28	0,965131033	0,034468965	0,000138382	48819
Environmental Science(all)	28	1,077970601	0,03849895	0,000340097	37374
Social Sciences(all)	28	0,759294713	0,027117668	0,000622247	30276
Immunology and Microbiology(all)	28	0,49360865	0,01762888	4,96455E-05	24850
Pharmacology, Toxicology and	28	0,535271939	0,019116855	5,14397E-05	23788
Chemical Engineering(all)	28	0,540318939	0,019297105	4,36273E-05	23009
Neuroscience(all)	28	0,335912388	0,011996871	4,49182E-05	20453
Arts and Humanities(all)	28	0,351748347	0,012562441	4,51971E-05	18687
Psychology(all)	28	0,263697498	0,009417768	4,32255E-05	13606
Business, Management and Accounting(all)	28	0,291108083	0,010396717	6,36367E-05	11906
Energy(all)	28	0,258782488	0,009242232	2,79859E-05	10802
Health Professions(all)	28	0,191870279	0,00685251	1,42973E-05	9191
Economics, Econometrics and Finance(all)	28	0,186481992	0,006660071	2,41876E-05	8242
Nursing(all)	28	0,108975686	0,003891989	5,70691E-06	6586
veterinary(all)	28	0,15396094	0,005498605	8,44767E-06	6232
Decision Sciences(all)	28	0,129124873	0,004611603	1,77136E-05	4999
Multidisciplinary	28	0,040400055	0,001442859	1,14805E-06	2992
Dentistry(all)	28	0,054457278	0,001944903	3,54747E-06	2990
					1244663

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1,22089	26	0,04696	115,354	4,7E-238	1,51084
Within Groups	0,29676	729	0,00041			
Total	1,51765	755				

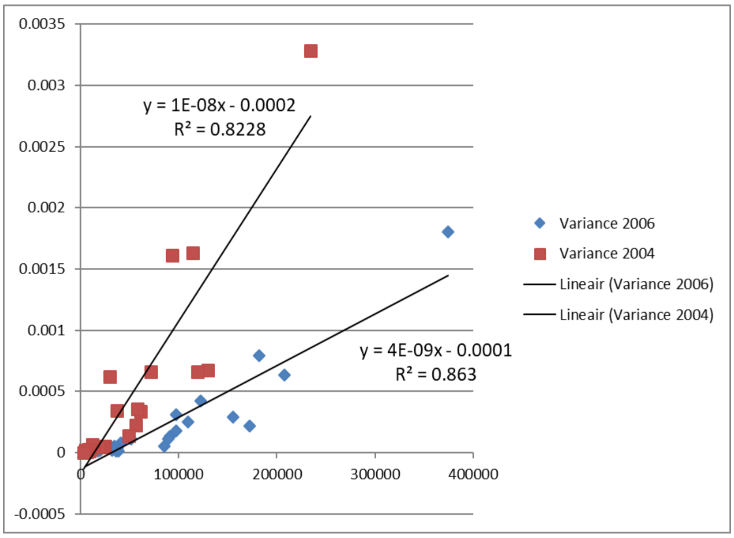
	Anova: Single Factor			2016	
Groups	*Count*	*Sum*	*Average*	*Variance*	*(Total output)*
Medicine(all)	28	4,176430818	0,149158244	0,001803115	373988
Engineering(all)	28	2,944519649	0,105161416	0,0006329	207832
Computer Science(all)	28	2,553668446	0,091202444	0,000792978	181897
Biochemistry, Genetics and Molecular Biology(all)	28	1,990931172	0,071104685	0,000220004	172660
Physics and Astronomy(all)	28	1,974952625	0,070534022	0,000289752	155115
Materials Science(all)	28	1,607755158	0,057419827	0,000419524	122512
Agricultural and Biological Sciences(all)	28	1,613323489	0,057618696	0,00024872	109862
Social Sciences(all)	28	1,400300485	0,050010732	0,000307619	97584
Mathematics(all)	28	1,329348765	0,047476742	0,00017757	97397
Chemistry(all)	28	1,119165924	0,039970212	0,000130252	90712
Environmental Science(all)	28	1,240739174	0,044312113	0,000112045	89431
Earth and Planetary Sciences(all)	28	0,960398597	0,03429995	5,00486E-05	85158
Arts and Humanities(all)	28	0,707528817	0,025268886	0,000112925	51969
Energy(all)	28	0,631851929	0,02256614	8,10507E-05	41393
Chemical Engineering(all)	28	0,462456109	0,01651629	1,51027E-05	38399
Neuroscience(all)	28	0,376289163	0,013438899	3,95646E-05	37574
Immunology and Microbiology(all)	28	0,429008484	0,015321732	1,41927E-05	36431
Business, Management and	28	0,495443546	0,017694412	4,96432E-05	34623
Pharmacology, Toxicology and	28	0,429875665	0,015352702	2,23971E-05	32558
Psychology(all)	28	0,34377275	0,012277598	3,61839E-05	32014
Economics, Econometrics and Finance(all)	28	0,277008092	0,009893146	2,19805E-05	20249
Health Professions(all)	28	0,217192679	0,007756881	1,03818E-05	17873
Multidisciplinary	28	0,196083781	0,007002992	1,45546E-05	15097
Nursing(all)	28	0,162714919	0,005811247	1,21895E-05	14776
Decision Sciences(all)	28	0,183925148	0,006568755	4,97458E-06	14410
veterinary(all)	28	0,112483235	0,004017258	7,53079E-06	7815
Dentistry(all)	28	0,062831382	0,002243978	3,86113E-06	5118
					2184447

ANOVA

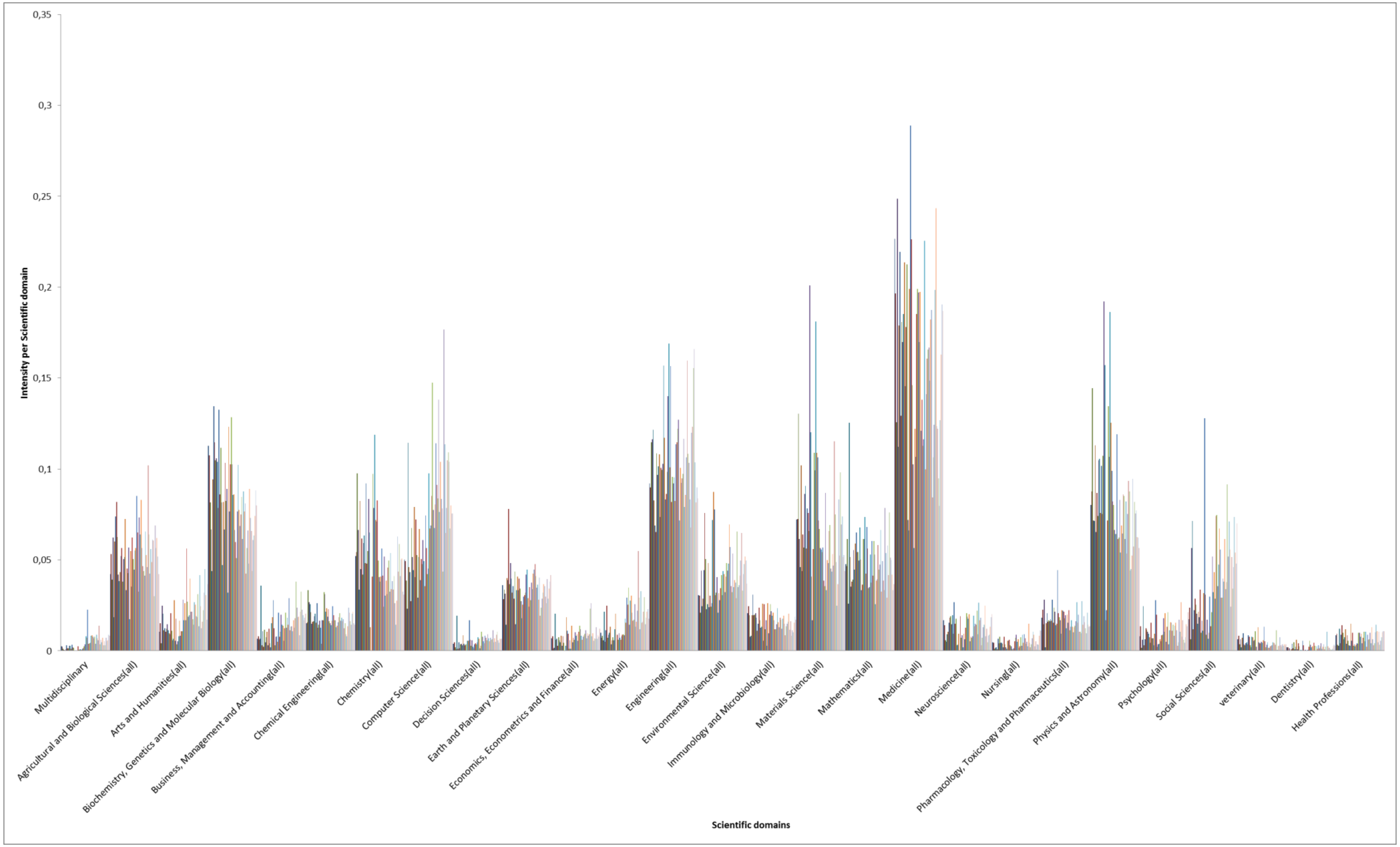
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0,93994	26	0,03615	173,34	0	1,51084
Within Groups	0,15204	729	0,00021			
Total	1,09198	755				

Table 2: The results to discuss the validity of hypothesis 3.

Table 2 presents the ANOVA output for an input of values representing the intensity in scientific output of the considered countries in the years 2004 and 2016. For both cases the differences in intensity are significant ($p<0.05$, and for 2016 = 0). This means that there is still variance among the different countries. In other words, in 2016 it was not the case that all science systems had the same distribution of the intensity within a scientific domain. Meanwhile the value of variance within groups had nearly halved. The first minor evidence for the existence of a "European Union" convergence club is given. Herewith the hypothesis that countries have moved into the same science domains is not rejected. A reasonable counter argument is that the variance within groups is still large and especially for the research areas wherein the Member States have the largest output – the science domains that characterise the convergence club. Therefore, a scatterplot (Graph 1) for the relation between the scientific output and the variance of both data selections is inserted. The calculated p-value for the difference between the two correlations is <0.0001. Concluding, although there is still variation of intensity, the difference between 2004 and 2016 is significant; hypothesis 3 is accepted.

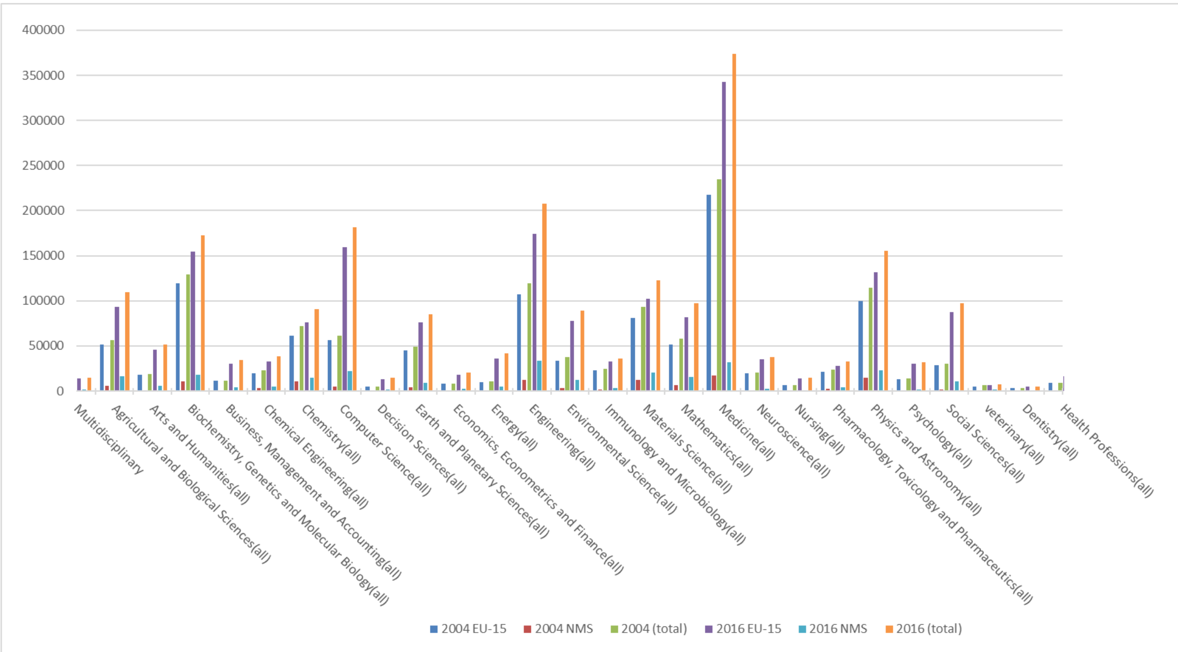


Graph 1: The relation between cumulated science output (X-axis) and variance (Y-axis).



Graph 2: Distribution of intensity of scientific output for all 28th Member States in the year 2004 (dark colours) and 2016 (light colours) .

The analysis of hypothesis 4 gives some context for the validation hypothesis 3. Graph 2 shows the intensity per science domain for each country in 2004 and 2016. The figure appears to be hard to interpret, but when a closer look is taken, it is evident that the output in certain areas not only increased, but also has become more dense. For example in case of “Biochemistry, Genetics and Molecular Biology” can be seen that in 2004 some gaps occur among the output. It means that the intensity of output in this science domain has been considerably lower regarding some countries. In 2016 these gaps have diminished, in other words, more countries have put their focus on this area; a larger percentage of their output is in the “Biochemistry, Genetics and Molecular Biology”. Remarkable is the increase of intensity in computer sciences. Regarding Graph 2, the first conclusion is that the science portfolio of the European Union is very diverse and still has a strong focus on medical and biology related sciences. Also, the total scientific output in “Engineering” is noticeable. To clarify how science systems from other convergence clubs may have contributed to focal point of knowledge generation in the European Union Graph 3 is relevant. It shows that initial output of the new Members States is negligible in comparison to the initial output of the EU-15 countries. In other words, the dispersion of the science portfolio of the European Union cannot be designated to the accession of new Member States. However, the hypothesis that the “European Union” science system will be more divers in 2016 is proven based on the presented graphs. Significantly is the growth in the earlier discussed computer sciences, but also the growth in “Arts and Humanities” and “Social Sciences”. The total values more than doubled when 2004 and 2016 are compared. This is also true for some smaller science domains, but the output in such domains is trivial compared to total output values.



Graph 3: The absolute scientific output for all Member States in different compositions for the years 2004 and 2016.

Discussion

In this section the results are interpreted and extrapolated by relating them to the theoretical frame work of this thesis. Beside this, limitations, considerations and relevance of this research will be further discussed.

Findings

To test hypothesis 1 and 2 the straightforward method established by Sala-i-Martin (1996) is used. Because of time limitation only β-convergence is calculated. The following step is not made; to test for σ-convergence among science system. As this value depends on the β-value[6], it will not affect the conclusions for the finding of *absolute convergence*. The first test confirms hypothesis 1; during the time period 2004 and 2016 convergence has taken place among the science systems of the European Union. Comparing this value to the β-value estimated by Horlings & van den Besselaar (2011) for the worldwide science output between 2000 and 2008 ($\beta = -0.008$), the pattern of convergence among the European Union countries is determined as strong ($\beta_{04\text{-}16}$: -0.017). Strong convergence among science systems means that smaller systems do grow considerably faster than larger ones. The fact that *absolute convergence* is found, made a proceeding study on *conditional convergence* relevant to further understanding what the main drivers are of the development of the science systems in the European Union.

Hypothesis 2 is not validated statistically as the difference between R-values of the functions is not significant. All other points made about hypothesis 2 should be viewed bearing this in mind. First of all, a stronger pattern of convergence is observed for the period of 2004-2016 compared to the period of 1996-2004. This is relevant regarding the ability of mainly new Member States[7] to generate sufficient knowledge. The hidden meaning of this point is the idea that "lagging" countries are not able to develop their science systems; increase their scientific output significantly. As shown with the proof for hypothesis 1, the smaller systems do grow faster than larger systems for the period 2004-2016. When taking a closer look at the data collection, these smaller systems are mainly new Member States. The outcomes when testing for ratification of hypothesis 2 do have some connotation. The results suggest that since their accession, new Member States are on their rise in respect of their science systems in comparison to the growth of larger science systems. As discussed in theory section, this increase can be explained by the interrelatedness between nations and their science systems. The assumption is that this is enforced because of accession to the European Union. In other words, "lagging countries" do catch up, however the question is if these countries are able to scale the infrastructure fast enough to produce the same scientific output as large science systems. To answer this question and how this problem can be overcome a very integrative approach is needed. To start with mapping the *conditional convergence* of science systems in the European Union.

This approach is also required to extend justifying the findings for the tests of hypothesis 3 and 4. It was not possible to test these two hypotheses with the same clarity as is done for the first two hypotheses. The output of the ANOVA to test if the Member States have become more similar in their scientific focus ratified hypothesis 3. Also, statistical prove was given by comparing the growth of total scientific output per domain and the change in variance among the intensity per science domain. Both outcomes support the argumentation that accession to the European Union stimulate countries that have been focussing on other science domains, to move to disciplines that are the focal point of other members of the "bigger system". This argument is aligned with the discussed dynamics of science systems and the occurrence of co-evolvement in path dependency of a country when a nation is

[6] σ-Convergence (takes place when $\sigma_{t+T} < \sigma_{t)}$ is calculated by comparing the standard deviations of the $\log(Y_t)$ for both time values (Horlings & van den Besselaar, 2011).

[7] Luxembourg is although it is one of the found8ing Member States, the country with the largest growth rate. On the other hand, regarding the size of its science system and the theory of convergence, the performance of this country is valid.

subjected to radical change. Notwithstanding the resemblance with the theory, in opinion of the researcher the results are to minimal to formulate a unambiguously statement about the acceptance of hypothesis 3. The confirmation within in the tests performed is evidently, however it cannot hardly be extrapolated. Therefore better comparative analysis should be performed. Some trial tests with factor analysis showed on the first sight interesting results. Unfortunately, it was not possible to perform this entire analysis. If such methods are considered to further investigate this topic, it is highly recommended to integrate also some demographical factors.

The fourth hypothesis states that the final (in 2016) “European Union” convergence club is more comprehensive than the initial one (2004). The science domains that characterise the scientific output in the European Union according to literature are medical and biology related sciences (Horlings &van den Besselaar, 2011). This has been also the case for the used dataset (SJR, 2016). However, the analysis in this research shows that science portfolio of the “convergence club” of the European Union in 2004 had also a serious output in other science domains, like engineering, material sciences and physics. All preceding science domains have larger outputs in 2016. “Computer sciences”, “Arts and Humanities” and “Social Sciences” (also the output in other social sciences like business and decision sciences) appear to be added to the “science portfolio” of the European Union. Horlings & van den Besselaar (2011) characterises “Computer Sciences” as a science domain of countries from the convergence club of “Emerging economies”, so most of the countries that accessioned the European Union in 2004. From this point view the underlying assumption of enlargement of the scientific portfolio through transmission of new knowledge is verified. But when the total output in this domain is further examined, the output of these new Member States is rather small compared to the output of the EU-15 countries. Therefore, a full confirmation of the assumption is not justified. Another support for the reconsideration of this assumption comes from the remarkable growth in “Arts and Humanities” and “Social Sciences”[8]. For both groups of countries (EU-15 countries and new Member States) these fields are minimal represented in 2004. Thus, the emerge in this field of science cannot be subscribed to the transmission of knowledge that originates from other convergence clubs. Therefore, hypothesis 4 is accepted, but the escorting assumption does not hold.

The research

To obtain the previously discussed findings, a quantitative study was performed. This was a logic step considering the entire dataset that was requested to form the input for this bachelor thesis. The choice to study this topic was a combination of personal interest of the researcher to investigate the European Union and the gap in literature that was found shortly after the decision was made for the geographical and political research domain. Driven by fascination for complex systems like the European Union, this study has been performed to the best ability of the researcher. The selection of articles appears to be minimal and less comparative. Unfortunately this could not be prevented because studying knowledge generation with this approach is rather new. Therefore the theory of path dependency is also discussed to provide some context for the occurrence of convergence in science systems. However, as mentioned above, the results invite for more (explanatory) research on this topic. Because this further extend was not possible, the outcomes and conclusion of this research should be considered and valued as indicative descriptive results.

[8] These science domains are not considered in the study of Horlings & van den Besselaar (2011).

Conclusion

The main research question of this study: "What convergence patterns in science systems can be identified among the Member States of the European Union?", led to two sub-research topics:

- The appearance of convergence considering of the national science systems. By taking into account the growth of the scientific output of 28th countries in the European Union.
- The status of the "European Union" convergence club for the last moment that was covered by the dataset (2016).

The conclusion for the research on the first topic is that absolute convergence has clearly occurred among the Member States. Thereby, the pattern of convergence appear to be stronger since the big enlargement of the European Union in 2004. This implies that the science system of the European Union is developing stronger since new countries entered. As this study only looked into absolute convergence, the drivers of this stronger convergence pattern are not understood. A consecutive study on conditional convergence for the same research topic is viewed as a way to procure this understanding.

The aim to determine the "European Union" convergence club has also led to the insight that actually not the countries that originate from other convergence clubs are responsible for the development of new characterising science domains across the EU. The "European Union" convergence club according to the scientific output in 2016 is characterised by: Medical and biology related sciences, engineering (all previous science domains are also focal point within the EU for 2004), computer science, social sciences, and arts and humanities. There is still room to investigate the precise relationships between these knowledge domains and how the exactly developed.

Summarising, absolute convergence have taken place among the Member States. And the convergence club of the European Union is divers. These findings invite for explanatory research on underlying conditions. This research should also be supported by the governing institutions of the European Union. Outcomes of such studies will guide better fitting policy making. Until then, European Commission can set an agenda on how the knowledge generation, especially in small systems, can keep growing. Small systems might be not able to grow as big as large science systems, due to for example infrastructural limitations.

A last word in regard of the Horizon 2020 programme. The focus points of the programme do match the characterising science domains of countries in the European Union quite well. Like medical sciences for the first (Excellent Sciences), the impressive new focus on computer science that was found in this study for the second (Leadership in Enabling and Industrial Technologies), and a combination guided by the developed focus in social sciences for the third (Societal Challenges). The last focus topic also speaks for further development in considered a very appropriate "science domain" in respect to the interrelatedness, (infrastructural) network and diversity of the EU science portfolio,. The study of interdisciplinary and highly complex problems. This may not bring the "European Union" in a leading position when it comes to innovation, but definitely the role which is crucial for innovation; providing context and a more comprehensive understanding of science.

Resources

Audretsch, D. B., Bönte, W., & Keilbach, M. (2008). Entrepreneurship capital and its impact on knowledge diffusion and economic performance. *Journal of business venturing*, *23*(6), 687-698.

Bellamy, R. (2017). A European republic of sovereign states: sovereignty, republicanism and the European Union. *European Journal of Political Theory*, *16*(2), 188-209.

Bonaccorsi, A. (2008). Search regimes and the industrial dynamics of science. *Minerva*, *46*(3), 285.

Bryman, A. (2012). *Social Research Methods* (4th ed.). Oxford, United Kingdom: Oxford University Press.

Curran, C. S., & Leker, J. (2011). Patent indicators for monitoring convergence–examples from NFF and ICT. *Technological Forecasting and Social Change*, *78*(2), 256-273.

Dedman, M. (2006). *The origins and development of the European Union 1945-1995: a history of European integration*. Routledge.

European Commission.(EC) (n.d.a). Knowledge Triangle and Innovation. Retrieved January 26, 2018, from Knowledge Triangle and Innovation. (n.d.). Retrieved 26th of January, 2018, from https://ec.europa.eu/education/policy/higher-education/knowledge-innovation-triangle_en

European Commission. (ECb) (n.d.b). Horizon 2020. Retrieved 26th of January, 2018, from http://ec.europa.eu/programmes/horizon2020/en/

European Union . (2017). Official website of the European Union.
Retrieved from: https://europa.eu/european-union/index_en
Last time visited: 8th of December 2017

Eurostat. (2017). Retrieved from: http://ec.europa.eu/eurostat/data/database
Last time visited: 8th of December 2017

Freeman, C., & Soete, L. (2009). Developing science, technology and innovation indicators: What we can learn from the past. *Research policy*, *38*(4), 583-589.

Hashimoto, T. & Rhimes, M. (2017). Reviewing European Union Accession: Unexpected Results, Spillover Effects, and Externalities. *Brill Nijhoff*. Leiden.

Hessels, L. K. (2013). Coordination in the science system: theoretical framework and a case study of an intermediary organization. *Minerva*, *51*(3), 317-339.

Hirooka, M. (2006). *Innovation dynamism and economic growth: A nonlinear perspective*. Edward Elgar Publishing.

Hirsch, P. M., & Gillespie, J. J. (2001). Unpacking path dependence: Differential valuations accorded history across disciplines. *Path dependence and creation*, 69-90.
Statistics Solutions. (n.d.). Comparing Correlation Coefficients. Retrieved 26th of January, 2018, from http://www.statisticssolutions.com/comparing-correlation-coefficients/

Horlings, E., & Van den Besselaar, P. (2011, September). Convergence in science: Growth and structure of worldwide scientific output, 1993–2008. In *Science and Innovation Policy, 2011 Atlanta Conference on* (pp. 1-19). IEEE.

Jeong, S., & Lee, S. (2015). What drives technology convergence? Exploring the influence of technological and resource allocation contexts. *Journal of Engineering and Technology Management*, *36*, 78-96.

Katz, J. S. (1999). The self-similar science system. *Research policy*, *28*(5), 501-517.

Leydesdorff, L. (2000). Is the European Union becoming a single publication system?. *Scientometrics*, *47*(2), 265-280.

Mahoney, J. (2000). Path dependence in historical sociology. *Theory and society*, *29*(4), 507-548.

Microsoft. (2017). Office Products.
Retrieved from: https://products.office.com/en/products?tab=O-Home
Last time visited: 8th of December 2017

Popadiuk, S., & Choo, C. W. (2006). Innovation and knowledge creation: How are these concepts related?. *International journal of information management*, *26*(4), 302-312.

Reillon, V. (2015, September). *Overview of EU Funds for research and innovation* [PDF]. European Union.

Quintane, E., Mitch Casselman, R., Sebastian Reiche, B., & Nylund, P. A. (2011). Innovation as a knowledge-based outcome. *Journal of Knowledge Management*, *15*(6), 928-947.

Scimago Journal & Country Rank (SJR). (2017).
Retrieved from: http://www.scimagojr.com/countryrank.php
Last time visited: 8th of December 2017

Solow, Robert M.; Dorfman, Robert; Samuelson, Paul (1958). *Linear programming and economic analysis*. New York: McGraw-Hill.

Stieglitz, N. (2004). Strategie und Wettbewerb in konvergierenden Märkten. Dt. Univ. *Verl., Wiesbaden*. In: Jeong, S., & Lee, S. (2015). What drives technology convergence? Exploring the influence of technological and resource allocation contexts. *Journal of Engineering and Technology Management*, *36*, 78-96.

Valdaliso, M., Magro, J., Navarro, E., Jose Aranguren, M., & Wilson, R. (2014). Path dependence in policies supporting smart specialisation strategies: Insights from the Basque case. *European Journal of Innovation Management*, *17*(4), 390-408.

Veuegelers, R. (2016). *The European Union's growing innovation divide*. Bruegel Commission.
PDF: http://bruegel.org/2016/04/the-european-unions-growing-innovation-divide/

Veugelers, R., & Mrak, M. (2009). The knowledge economy and catching-up member states of the European Union. *European Commission, http://ec. europa. eu/invest-in-research/pdf/download_en/kfg_report_no5. pdf*.

Warleigh-Lack, A. (2008). *European Union: the basics*. Routledge.

Weller, C. (17, June 15). The 16 most innovative countries in the world. *Business Insiders*. Retrieved 26th of January, 2018, from https://www.businessinsider.nl/most-innovative-countries-in-the-world-2017-6/?international=true&r=US